Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 183 (2011) 92–100
FUZZY sets and systems
www.elsevier.com/locate/fss

# Interval-valued probability density estimation based on quasi-continuous histograms: Proof of the conjecture

J.-F. Crouzet[a], O. Strauss[b,*]

[a] I3M Université Montpellier II, Place Eugéne Bataillon, 34095 Montpellier Cedex, France
[b] LIRMM Université Montpellier II, 161, rue Ada, 34392 Montpellier Cedex 5, France

Received 17 June 2010; received in revised form 15 February 2011; accepted 27 February 2011
Available online 4 March 2011

**Abstract**

The sensitivity of histogram computation to the choice of a reference interval and number of bins can be attenuated by replacing the crisp partition on which the histogram is built by a fuzzy partition. This involves replacing the crisp counting process by a distributed (weighted) voting process. The counterpart to this low sensitivity is some confusion in the count values: a value of 10 in the accumulator associated with a bin can mean 10 observations in the bin or 40 observations near the bin. This confusion can bias the statistical decision process based on such a histogram. In a recent paper, we proposed a method that links the probability measure associated with any subset of the reference interval with the accumulator values of a fuzzy partition-based histogram. The method consists of transferring counts associated with each bin proportionally to its interaction with the considered subset. Two methods have been proposed which are called precise and imprecise pignistic transfer. Imprecise pignistic transfer accounts for the interactivity of two consecutive cells in order to propagate, in the estimated probability measure, counting confusion due to fuzzy granulation. Imprecise pignistic transfer has been conjectured to include precise pignistic transfer. The present article proposes a proof of this conjecture.


*Keywords:* Quasi-continuous histograms; Imprecise probability; Capacity

## 1. Introduction

Although substantial research has been focused on providing scientists with graphical tools for displaying the probability density function (pdf) underlying a set of observations, the histogram is the oldest and still most widely used method [11]. The major argument in favor of histogram-based analysis is its easy computation and interpretation. Building a histogram on the basis of real observations involves partitioning a reference interval $\Omega$ of the real line into $p$ cells (or bins) $(C_k)_{k=1\ldots p}$ and counting the number of observations in agreement with each cell. When the number of observations tends to infinity, the ratio of the count associated with each cell and the number of observations tends to the probability of the bin based on the pdf underlying the observation process [12,10].

* Corresponding author. Tel.: +33 467 418 587; fax: +33 467 418 500.
*E-mail addresses:* Jean-Francois.Crouzet@univ-montp2.fr (J-F. Crouzet), strauss@lirmm.fr (O. Strauss).

0165-0114/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.fss.2011.02.019

A histogram can therefore be viewed as a rough approximate decomposition of the sought after pdf on the basis of the cells, with the roughness of this decomposition being due to both the cell widths and the number of observations. The histogram – thanks to its close relation with the pdf – is an instrumental tool for designing discrete statistical analysis methods including modes, moments and quantile estimations [15] or comparison of two sets of observations (e.g. in image processing [1,3]). Such methods are usually based on representing the link between the histogram and the pdf by means of an interpolation function to obtain an estimate of the pdf at any location of $\Omega$. As shown in [13], interpolating a histogram with a translation invariant kernel leads to defining a method for estimating the pdf that is equivalent to the Parzen–Rosenblatt method, using a kernel that is not translation invariant. Using a histogram instead of Parzen–Rosenblatt density estimation to define statistical tools has the advantage of leading to low-computational algorithms.

However, the roughness of the histogram representation of the pdf and its sensitivity to the cell width and position considerably lower the robustness of histogram-based statistical tools. Different proposals have been put forward in the relevant literature to lower this sensitivity (see e.g. [6] and references therein), including replacement of the crisp partition on which the histogram is built by a fuzzy partition[1] [5]. But this sensitivity still exists. It would thus be interesting to have a tool for evaluating the sensitivity of a histogram-based statistical method. It would also be nice to have a framework that makes it possible to define discrete statistical tools that are robust to the partitioning of $\Omega$.

This is the core idea of the quasi-continuous histogram (qch) framework proposed in [13]. The qch framework allows precise and imprecise estimations of the probability of any interval $W$ of $\Omega$ based on the pdf underlying the considered set of observations, and thus the histogram. It is based on transferring the counts $a_k$ associated with each fuzzy cell $C_k$ to the considered interval, proportionally to the interaction between the cell and the considered interval. Three methods have been proposed.

The first one, called pignistic transfer, is closely related to Parzen–Rosenblatt kernel estimation. Within this method, the interaction between a cell $C_k$ and an interval $W$ is defined as being the ratio[2] $|W \cap C_k|/|C_k|$. The number of counts associated with $W$, based on this transfer, is defined by

$$\hat{Nb}(W) = \sum_{k=1}^{p} \frac{|W \cap C_k|}{|C_k|} a_k.$$

Then, the probability of $W$ is simply obtained by dividing $\hat{Nb}(W)$ by the number of observations. It is the most natural method for interpolating histograms and leads to a precise-valued probability $\hat{P}(W)$.

The second one, called possibilistic transfer, leads to an interval-valued probability. Within this method, the interaction between a cell $C_k$ and the interval $W$ is defined by two dual measures: the possibility and necessity of $W$ restricted to $C_k$. This transfer provides an interval-valued number of counts $[\underline{Nb}(W), \overline{Nb}(W)]$ associated with $W$ defined by

$$\underline{Nb}(W) = \sum_{k=1}^{p} N_k(W) a_k, \quad \overline{Nb}(W) = \sum_{k=1}^{p} \Pi_k(W) a_k,$$

where $\Pi_k(W)$ and $N_k(W)$ are the possibility and necessity of $W$ restricted to $C_k$. Then, the interval-valued probability $[\underline{P}(W), \overline{P}(W)]$ of $W$ is simply obtained by dividing the interval $[\underline{Nb}(W), \overline{Nb}(W)]$ by the number of observations. It has been proved that the possibilistic-based estimate includes the pignistic-based estimate due to the known domination properties between summative and maxitive kernels [7]. Moreover, the construction of this interval-valued probability makes it the convex hull of all probabilities of $W$ that should have been obtained by interpolating the histogram by a set of kernel smoothing methods. If the partition is crisp, the interval-valued probability remains imprecise.

The third one, called imprecise pignistic transfer, leads to an interval-valued probability which is more specific than the previous one. Within this method, the interaction between a cell $C_k$ and interval $W$ is defined by two dual particular capacities derived from the interaction measure used in pignistic transfer. The idea underlying this method is as follows. Due to fuzzy partitioning, interpretation of counts associated with each cell becomes ambiguous. In fact, a count value of 6 associated with a cell can denote six observations fully belonging to the cell but also 20 observations belonging to the cell at a level of 0.3. This method aims at transferring the counting imprecision due to fuzzy partitioning estimation

[1] When doing this replacement, the number of cells is increased to ($p$+1).

[2] The measure $|A|$ when $A$ is a fuzzy set will be defined in the next section.

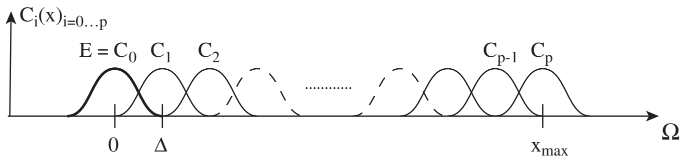


Fig. 1. Fuzzy partition of $\Omega$.

of the probability of interval $W$. In fact, if the partition is crisp, then the count values are no longer ambiguous and thus the imprecise pignistic-based estimate is precise. Numerous examples show that an imprecise pignistic-based estimate always includes a pignistic-based estimate. This property has thus been conjectured [13, Conjecture 2].

This paper aims at proving this conjecture. The framework and notations are presented in Section 2. Section 3 is dedicated to the proof itself.

## 2. Framework and notations

Let $\Omega$ be a reference interval of the real line. We can, without any loss of generality, assume that $\Omega = [0, x_{max}]$. Let $(C_k)_{k=0\ldots p}$ be a strong[3] uniform[4] fuzzy partition of $\Omega$ [9] generated by a generic, symmetric, unimodal, centered and bounded fuzzy subset $E$ whose membership function is $\mu_E$ (see e.g. [6]). That is, there are $(p+1)$ values $m_k$ of $\Omega$ such that $\forall k \in \{0 \ldots p\}$, $\forall x \in \Omega$, $\mu_{C_k}(x) = \mu_E(x - m_k)$ (see Fig. 1). The $(p+1)$ values $(m_k)_{k=0\ldots p}$ are called modes of the partition. Let $|A|$ denote the granularity [7] of a fuzzy subset $A \subset \Omega$ defined by $|A| = \int_\Omega \mu_A$, with $\int$ being the Lebesgue integral. Let $\Delta$ be the granularity of $E$: $\Delta = |E| = \int_\Omega \mu_E$. We consider, as usual, partitions such that $\forall x \in \Omega$, there are at most two consecutive cells, $C_k$ and $C_{k+1}$, such that $\mu_{C_k}(x) > 0$ and $\mu_{C_{k+1}}(x) > 0$. Let $\cap$ denote the intersection of fuzzy sets based on the minimum t-norm and $\cup$ denote the union of fuzzy sets based on the maximum t-conorm.

The symmetry of $E$ coupled with the proposed definition of the partition induces two particular properties on $E$. First, the length of the support of $E$ (i.e. its 0-level cut) equals twice its granularity. Second, $\forall x \in [0, \Delta/2]$, $\mu_E(x) = 1 - \mu_E(\Delta - x)$.

Let $(x_i)_{i=1\ldots n}$ be $n$ observations belonging to $\Omega$. Building a histogram of these observations consists of associating a value $a_k$ (called an accumulator) with each (fuzzy) cell $C_k$ defined by $a_k = \sum_{i=1}^n \mu_{C_k}(x_i)$. If the cell $C_k$ is crisp, then the value $P_k = (a_k/n)$ can be seen as an empirical estimate of its probability based on the set of observations. When the cell $C_k$ is fuzzy, then this value $P_k = (a_k/n)$ is the Zadeh extension of its empirical probability [16].

Let $W = [\underline{x}, \overline{x}] \subset \Omega$ be an interval of $\Omega$. In [13], it is proposed to estimate the probability of $W$ by a kind of interpolation based on transferring the $(P_k)_{k=0\ldots p}$ to W by

$$\hat{P}(W) = \sum_{k=0}^{p} P_k.\rho_W(k), \tag{1}$$

for the precise pignistic transfer, with

$$\rho_W(k) = \frac{|C_k \cap W|}{|W|} = \frac{1}{|W|}\int_W \mu_{C_k},$$

and by

$$\overline{P}(W) = \sum_{k=0}^{p} P_{(k)}(\nu_W(A_{(k)}) - \nu_W(A_{(k+1)})), \tag{2}$$

$$\underline{P}(W) = \sum_{k=0}^{p} P_{(k)}(\nu^c_W(A_{(k)}) - \nu^c_W(A_{(k+1)})), \tag{3}$$

[3] Strong means that $\forall x \in \Omega$, $\sum_{k=0}^{p} \mu_{C_k}(x) = 1$.

[4] Uniform means that every cell of the partition can be obtained by translating a single fuzzy subset.

for the imprecise pignistic transfer, where $(\cdot)$ indicates a permutation such that $P_{(0)} \leq \cdots \leq P_{(p)}$, and $\{A_{(k)}\}_{k=0\ldots p}$ are subsets of $\Theta = \{0 \ldots p\}$ such that $A_{(k)} = \{(k), \ldots, (p)\}$, and $A_{(p+1)} = \emptyset$. $v_W$ is a capacity defined by

$$\forall A \subseteq \Theta, \quad v_W(A) = \frac{|(\bigcup_{k\in A} C_k) \cap W|}{|(\bigcup_{k\in\Theta} C_k) \cap W|} = \frac{\int_W \sup_{k\in A} \mu_{C_k}}{\int_W \sup_{k\in\Theta} \mu_{C_k}}, \tag{4}$$

and $v_W^c$ its dual capacity defined by $\forall A \subseteq \Theta$, $v_W^c(A) = 1 - v_W(A^c)$, with $A^c$ being the complementary set of $A$ in $\Theta$.

Eqs. (2) and (3) are simply discrete Choquet integrals of $\{P_k\}_{k=0\ldots p}$ with respect to the capacities $v_W$ and $v_W^c$. Eq. (1) can also be considered as a Choquet integral by considering the capacity $\phi_W$ defined by $\forall A \subseteq \Theta$, $\phi_W(A) = \sum_{k\in A} \rho_W(k)$.

In [13], the following conjecture has been proposed:

**Conjecture.** *The precise-valued probability $\hat{P}(W)$ defined by Eq.* (1) *always belongs to the interval-valued probability* $[\underline{P}(W), \overline{P}(W)]$ *defined by Eqs.* (2) *and* (3).

The proof of this conjecture will follow. It is based on proving that $\forall A \subseteq \Theta$, $v_W(A) \geq \phi_W(A)$. This domination property is sufficient to prove the conjecture (see e.g. [2]).

## 3. Proof of the conjecture

We show the inequality $\forall A \subseteq \Theta$, $v_W(A) \geq \phi_W(A) = \sum_{k\in A} \rho_W(k)$ by a constructive process: starting from the empty set of $\Theta$, we will incrementally construct the set $A$ by adding intermediate subsets of $A$ and prove that the inequality is not modified during this constructive process.

Let us first denote $\omega_\Theta = \sup_{k\in\Theta} \mu_{C_k}$, $\omega_A = \sup_{k\in A} \mu_{C_k}$ and $\eta_A = \sum_{k\in A} \mu_{C_k}$ (see Figs. 2–4). The conjecture can be written as follows:

$$|W| \int_W \omega_A \geq \int_W \omega_\Theta \int_W \eta_A. \tag{5}$$

Note that $|W| = \int_W \eta_\Theta = \int_W 1$. Inequality (5) is obvious if $A = \emptyset$.

For an easier development of the proof, we will also use another expression of the inequality (5):

$$\int_W (1 - \omega_\Theta) \int_W \omega_A - \int_W \omega_\Theta \int_W (\eta_A - \omega_A) \geq 0. \tag{6}$$

To build the subset $A$, we will consider two kinds of elements of $A$: isolated elements, i.e. values $k \in A$ such that $(k+1) \notin A$ and $(k-1) \notin A$ – which are associated with isolated cells of $\{C_k\}_{k\in A}$ – and those which are not isolated,

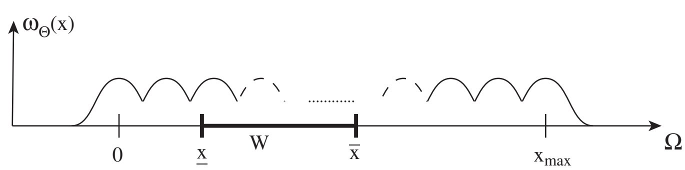


Fig. 2. Function $\omega_\Theta$ and interval $W$.

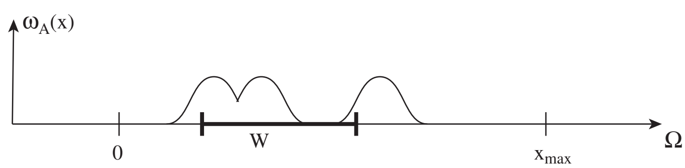


Fig. 3. Function $\omega_A$, with $A$ being composed of three elements of $\Theta$.

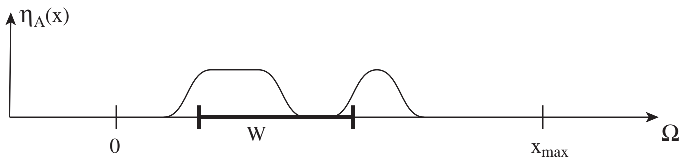


Fig. 4. Function $\eta_A$, with $A$ being composed of three elements of $\Theta$.

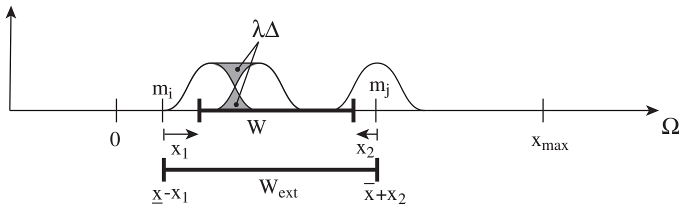


Fig. 5. Values $x_1$, $x_2$, $\lambda$ and $W_{ext}$.

i.e. values $k \in A$ such that $(k+1) \in A$ or $(k-1) \in A$ – which are associated with overlapping cells of $\{C_k\}_{k\in A}$. A group of consecutive non-isolated values of $A$ will be called a connected block. Of course, we are only considering elements in $A$ having a non-empty intersection with $W$, since the integral of other elements with respect to $W$ will equal zero. For simplicity, we will now consider $A$ as a subset of $\Theta$ or as the group of cells $\{C_k\}_{k\in A}$, i.e. as a collection of fuzzy subsets of $\Omega$.

We first need to define $W_{ext}$ as the smallest interval $[m_i, m_j]_{(i,j)\in\Theta, i\le j}$ that includes $W$ (see Fig. 5). Its length is $|W_{ext}| = n\varDelta = |W| + x_1 + x_2$, where $x_1, x_2 \in [0, \varDelta[$[5] (see Fig. 5), and $n = j - i$. We also define a parameter $\lambda$ by considering the fuzzy subset obtained by intersecting two consecutive cells of the partition. Clearly, the granularity of this intersection depends on both the shape of the generic subset $E$ and its granularity $\varDelta$. Let $\lambda\varDelta$ be the granularity of this intersection (see Fig. 5). Due to the symmetry of the generic function $\mu_E$, $0 \le \lambda \le \frac{1}{2}$ (0 if $E$ is crisp and $\frac{1}{4}$ if $E$ is a triangular fuzzy number).

The first step of the construction of $A$ consists of making subset $A'$ of all isolated elements of $A$. Since the cells composing $\{C_k\}_{k\in A'}$ are isolated, by construction $\int_W \omega_{A'} = \int_W \eta_{A'}$. Now, considering the fact that $\forall x \in \Omega$, $\omega_\Theta(x) \le 1$, thus $\int_W \omega_\Theta \le \int_W 1 = |W|$, and therefore $|W| \int_W \omega_{A'} \ge \int_W \omega_\Theta \int_W \eta_{A'}$. Thus $A'$ verifies the property (5).

The second step of the construction of $A$, and therefore of the proof of the conjecture, consists of adding each block of overlapping cells of $A$ one by one. We will prove that, when adding a block, property (5) is kept. Two cases will be considered depending on whether the considered block intersects interval $W$ or not. Let $B$ be this block ($B \subset A$) and $k$ its cardinality. Let $A' \subset A$ be a subset of $A$ such that $A' \cap B = \emptyset$. We will prove that if $A'$ verifies the inequality (5), then $A'' = A' \cup B$ also verifies this inequality.

In inequality (6), the two integrals $\int_W \omega_\Theta$ and $\int_W (1-\omega_\Theta)$ depend neither on the position nor on the size of $B$. Those two integrals are given by

$$\int_W \omega_\Theta = \int_0^{n\varDelta} \omega_\Theta - \int_0^{x_1} \omega_\Theta - \int_0^{x_2} \omega_\Theta = n(1-\lambda)\varDelta - \left(\int_0^{x_1} \omega_\Theta + \int_0^{x_2} \omega_\Theta\right),$$

$$\int_W (1-\omega_\Theta) = n\lambda\varDelta - \left(\int_0^{x_1} (1-\omega_\Theta) + \int_0^{x_2} (1-\omega_\Theta)\right).$$

We now have to consider two cases.

*First case*: $B$ is included in $W$.

[5] $[\cdot, \cdot[$ denotes a left-closed and right-open interval.

In that case, adding $B$ to $A'$ will increase $\int_W \eta_{A'}$ of the quantity $k\Delta$ (due to the symmetry of the generic function $\mu_E$) and increase $\int_W \omega_{A'}$ of the quantity $k\Delta - (k-1)\lambda\Delta = k(1-\lambda)\Delta + \lambda\Delta$. Thus, the inequality (6) will be preserved if and only if we show that

$$\left(n\lambda\Delta - \left(\int_0^{x_1}(1-\omega_\Theta) + \int_0^{x_2}(1-\omega_\Theta)\right)\right)(k(1-\lambda)\Delta + \lambda\Delta)$$
$$-\left(n(1-\lambda)\Delta - \left(\int_0^{x_1}\omega_\Theta + \int_0^{x_2}\omega_\Theta\right)\right)(k\lambda\Delta - \lambda\Delta) \geq 0$$

which can be simplified to

$$n\lambda\Delta + k\left[\left\{-\int_0^{x_1}(1-\omega_\Theta) + \lambda x_1\right\} + \left\{-\int_0^{x_2}(1-\omega_\Theta) + \lambda x_2\right\}\right] - \lambda(x_1+x_2) \geq 0.$$

Let us first suppose that $n \leq k+1$. This implies, by construction, that $n = k+1$ and thus $x_1 = x_2 = 0$. This is a trivial case: the inequality is preserved. Now let us suppose $n \geq k+2$. By replacing $n$ by $k+2$ in the previous equation, we obtain a new sufficient condition that can be written as follows:

$$k\left[\left\{-\int_0^{x_1}(1-\omega_\Theta) + \lambda x_1 + \lambda\frac{\Delta}{2}\right\} + \left\{-\int_0^{x_2}(1-\omega_\Theta) + \lambda x_2 + \lambda\frac{\Delta}{2}\right\}\right] + \lambda(2\Delta - (x_1+x_2)) \geq 0.$$

Consider $\varphi(x) := -\int_0^x(1-\omega_\Theta) + \lambda x + \lambda\Delta/2$, for $x \in [0, \Delta[$. As $\int_0^{\Delta/2}(1-\omega_\Theta) = \int_{\Delta/2}^{\Delta}(1-\omega_\Theta) = \lambda\Delta/2$, we get $\forall x \in [0, \Delta/2[$, $\varphi(x) \geq -\lambda(\Delta/2) + \lambda x + \lambda(\Delta/2) = \lambda x \geq 0$, and $\forall x \in [\Delta/2, \Delta[$, $\varphi(x) \geq -\lambda\Delta + \lambda x + \lambda(\Delta/2) = \lambda(x - \Delta/2) \geq 0$. Thus, as the sufficient condition is verified, subset $A'' = A' \cup B$ verifies the inequality.

*Second case*: $B$ is not included in $W$.

In this second part of the proof, we will first suppose that $n \geq k+2$, i.e. block $B$ only intersects one side of $W$ ($\underline{x}$ or $\overline{x}$) (the case $n \leq k+1$ will be treated at the end of this second part). We can, without any loss of generality, suppose that $B$ intersects the left part of $W$. The proof is based on computing how the different quantities involved in inequality (6) are modified for $W_{ext}$, then correcting those values with a quantity that will, in this case, depend only on $x_1$.

In the following, we differentiate two cases corresponding to the manner in which $B$ intersects the left side of $W$:

(a) *Block B is included in* $W_{ext}$: This case is fully illustrated by Fig. 5.

It can be easily established that adding $B$ to $A'$ will increase $\int_W \eta_{A'}$ of the quantity $k\Delta - \int_{\underline{x}-x_1}^{\underline{x}} \eta_B$ and increase $\int_W \omega_{A'}$ of the quantity $k(1-\lambda)\Delta + \lambda\Delta - \int_{\underline{x}-x_1}^{\underline{x}} \omega_B$. Thus, inequality (6) will be preserved if and only if we show that

$$\left(n\lambda\Delta - \left\{\int_0^{x_1}(1-\omega_\Theta) + \int_0^{x_2}(1-\omega_\Theta)\right\}\right)\left(k(1-\lambda)\Delta + \lambda\Delta - \int_{\underline{x}-x_1}^{\underline{x}}\omega_B\right)$$
$$-\left(n(1-\lambda)\Delta - \left\{\int_0^{x_1}\omega_\Theta + \int_0^{x_2}\omega_\Theta\right\}\right)\left(k\lambda\Delta - \lambda\Delta - \int_{\underline{x}-x_1}^{\underline{x}}(\eta_B - \omega_B)\right) \geq 0,$$

that can be rewritten as follows:

$$n\Delta\left[\lambda\Delta - \lambda\int_{\underline{x}-x_1}^{\underline{x}}\omega_B + (1-\lambda)\int_{\underline{x}-x_1}^{\underline{x}}(\eta_B - \omega_B)\right]$$
$$-k\Delta\left[\left\{-\int_0^{x_1}(1-\omega_\Theta) + \lambda x_1\right\} + \left\{-\int_0^{x_2}(1-\omega_\Theta) + \lambda x_2\right\}\right]$$
$$+\left\{\int_0^{x_1}(1-\omega_\Theta) + \int_0^{x_2}(1-\omega_\Theta)\right\}\int_{\underline{x}-x_1}^{\underline{x}}\omega_B$$
$$-\left\{\int_0^{x_1}\omega_\Theta + \int_0^{x_2}\omega_\Theta\right\}\int_{\underline{x}-x_1}^{\underline{x}}(\eta_B - \omega_B) - \lambda\Delta(x_1+x_2) \geq 0.$$

Since $n \geq k+2$ and $k \geq 2$, by definition, a sufficient condition is, after computation:

$$(k-2)\Delta\gamma + R \geq 0,$$

where

$$\gamma = \varphi(x_1) + \varphi(x_2) - \lambda \int_{\underline{x}-x_1}^{\underline{x}} \omega_B + (1-\lambda) \int_{\underline{x}-x_1}^{\underline{x}} (\eta_B - \omega_B)$$

and

$$R = 2\Delta(\varphi(x_1)+\varphi(x_2)) + \left\{ \int_0^{x_1} (1-\omega_\Theta) + \int_0^{x_2} (1-\omega_\Theta) - 4\lambda\Delta \right\} \int_{\underline{x}-x_1}^{\underline{x}} \omega_B$$
$$+ \left\{ 4(1-\lambda)\Delta - \left( \int_0^{x_1} \omega_\Theta + \int_0^{x_2} \omega_\Theta \right) \right\} \int_{\underline{x}-x_1}^{\underline{x}} (\eta_B - \omega_B) + \lambda\Delta(2\Delta - (x_1+x_2))$$

(still with $\varphi(x) := -\int_0^x (1-\omega_\Theta) + \lambda x + \lambda(\Delta/2) \geq 0$).

Two sub-cases that change the encountered integral increments have to be distinguished: $x_1 \in [0, \Delta/2[$ and $x_1 \in [\Delta/2, \Delta[$.

(i) $x_1 \in [0, \Delta/2[$: As $\int_{\underline{x}-x_1}^{\underline{x}} \omega_B = \int_0^{x_1} \omega_B(\underline{x} - x_1 + t)\,dt = \int_0^{x_1} (1-\omega_\Theta)$ we derive $\gamma = \varphi(x_1) + \varphi(x_2) - \lambda \int_0^{x_1} (1-\omega_\Theta) \geq -(1+\lambda)\int_0^{x_1}(1-\omega_\Theta) + \lambda x_1 + \lambda(\Delta/2) := f(x_1)$. A quick study of *f* shows that it increases from the value $\lambda(\Delta/2)$ to its maximum, and then decreases to the value $(1-\lambda)\lambda(\Delta/2) \geq 0$, such that $\gamma \geq 0$.

We have $R \geq 2\Delta\varphi(x_2) + \{\int_0^{x_2}(1-\omega_\Theta) - 4\lambda\Delta\}\int_0^{x_1}(1-\omega_\Theta) + \lambda\Delta(2\Delta-(x_1+x_2)) := g(x_1)$. Recalling that $\int_0^{\Delta}(1-\omega_\Theta) = \lambda\Delta$, we get

$$g'(x_1) = (1-\omega_\Theta(x_1)) \left\{ \int_0^{x_2} (1-\omega_\Theta) - 4\lambda\Delta \right\} - \lambda\Delta < 0,$$

and have to prove that $g(\Delta/2) \geq 0$. We thus compute

$$g\left(\frac{\Delta}{2}\right) = \lambda\frac{\Delta}{2} \left\{ \int_0^{x_2} (1-\omega_\Theta) - 2x_2 + 3\Delta - 4\lambda\Delta \right\} + 2\Delta\varphi(x_2) := h(x_2).$$

For $x_2 \in [0, \Delta/2[$, $g(\Delta/2) \geq \lambda\Delta^2(1-2\lambda) \geq 0$ (since $\lambda < \frac{1}{2}$). For $x_2 \in [\Delta/2, \Delta[$, we get after computation: $g(\Delta/2) \geq \lambda\Delta(x_2 + \Delta/2 - \frac{7}{4}\lambda\Delta) \geq \frac{1}{8}\lambda\Delta^2 \geq 0$. This completes point (i).

(ii) $x_1 \in [\Delta/2, \Delta[$: As $\int_{\underline{x}-x_1}^{\underline{x}} \eta_B = \int_{\underline{x}-x_1}^{\underline{x}} \omega_B = \int_0^{x_1} \omega_B(\underline{x} - x_1 + t)\,dt = \int_0^{x_1} \omega_\Theta = \lambda(\Delta/2) + \int_{\Delta/2}^{x_1} \omega_\Theta$ we get $\gamma \geq \varphi(x_1) - \lambda^2(\Delta/2) - \lambda\int_{\Delta/2}^{x_1} \omega_\Theta := f(x_1)$. As $f'(x_1) = (1-\lambda)(\omega_\Theta - 1) \leq 0$ and $f(\Delta) = \lambda(\Delta/2)(1-\lambda-(1-\lambda)) = 0$, we have $\gamma \geq 0$.

Now $R \geq 2\Delta(\varphi(x_1)+\varphi(x_2)) + \{\lambda(\Delta/2) + \int_{\Delta/2}^{x_1}\omega_\Theta\}\{\lambda(\Delta/2) + \int_0^{x_2}(1-\omega_\Theta) - 4\lambda\Delta\} + \lambda\Delta(2\Delta - (x_1+x_2)) := g(x_1)$. We compute

$$g'(x_1) = \omega_\Theta(x_1) \left\{ \int_0^{x_2} (1-\omega_\Theta) + 2\Delta - \frac{7}{2}\lambda\Delta \right\} - 2\Delta + \lambda\Delta \leq \int_0^{x_2} (1-\omega_\Theta) - \frac{5}{2}\lambda\Delta < 0$$

and

$$g(\Delta) = \Delta \left\{ -\frac{3}{2} \int_0^{x_2} (1-\omega_\Theta) + \lambda x_2 + \frac{5}{4}\lambda\Delta \right\}.$$

For $x_2 \in [0, \Delta/2[$, $g(\Delta) \geq \lambda(x_2 + \Delta/2) \geq 0$.

For $x_2 \in [\Delta/2, \Delta[$, $g(\Delta) \geq \lambda(x_2 - \Delta/4) \geq 0$.

This completes point (ii).

(b) *Block B is not included in $W_{ext}$.*

This case is fully illustrated by Fig. 6.

It can now be established that adding $B$ to $A'$ will increase $\int_W \eta_{A'}$ of the quantity $k\Delta - (\Delta/2 + \int_{\underline{x}-x_1}^{\underline{x}} \eta_B)$ and increase $\int_W \omega_{A'}$ of the quantity $k(1-\lambda)\Delta + \lambda\Delta - (\Delta/2 + \int_{\underline{x}-x_1}^{\underline{x}} \omega_B)$.

As $\forall t \in [\underline{x} - x_1, \underline{x}], \eta_B(t) = 1$ and $\omega_B(t) = \omega_\Theta$, we now get $\int_{\underline{x}-x_1}^{\underline{x}} \omega_B = \int_0^{x_1} \omega_\Theta$ and $\int_{\underline{x}-x_1}^{\underline{x}} (\eta_B - \omega_B) = \int_0^{x_1}(1-\omega_\Theta)$.

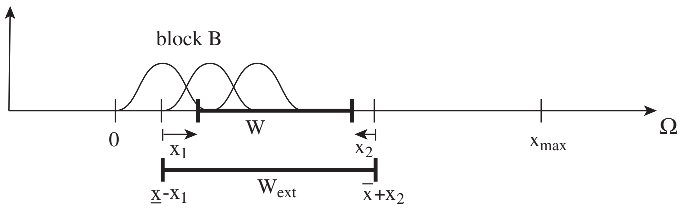


Fig. 6. Block $B$ is not included in $W_{ext}$.

Following the same reasoning as in case (a), a sufficient condition is also

$$(k-2)\Delta\gamma + R \geq 0.$$

Still considering $\varphi(x) := -\int_0^x (1-\omega_\Theta) + \lambda x + \lambda(\varDelta/2) \geq 0$, the computation yields $\gamma = \varphi(x_1) + \varphi(x_2) - \lambda \int_0^{x_1} \omega_\Theta + (1-\lambda)\int_0^{x_1}(1-\omega_\Theta) - \lambda(\varDelta/2) = \varphi(x_2) + \lambda(\varDelta/2) - \lambda(\varDelta/2) = \varphi(x_2) \geq 0$.

Let us now conclude the proof in the case $n \geq k+2$ by showing that $R \geq 0$ for all values of $x_1$ and $x_2$.

The expression of $R$ is as follows:

$$R = 2\varDelta\left(-\int_0^{x_2}(1-\omega_\Theta) + \lambda x_2 + \lambda\frac{\varDelta}{2}\right) + 2\varDelta\left(\int_0^{x_1}(1-\omega_\Theta) - \lambda x_1 + \lambda\frac{\varDelta}{2}\right)$$
$$+\int_0^{x_2}(1-\omega_\Theta)\left(\frac{\varDelta}{2} + \int_0^{x_1}\omega_\Theta\right)\int_0^{x_1}(1-\omega_\Theta)\left(\frac{\varDelta}{2} - \int_0^{x_2}\omega_\Theta\right) - \lambda\varDelta(x_1+x_2).$$

Let us denote $R := g(x_2)$. After simplification, we get the derivative $g'(x_2) = (1-\omega_\Theta(x_2))(\int_0^{x_1}\omega_\Theta - \frac{3}{2}\varDelta) - \omega_\Theta(x_2)\int_0^{x_1}(1-\omega_\Theta) - \lambda\varDelta \leq 0$. It therefore now must be shown that $g(\varDelta) \geq 0$. The calculus yields the following value, $g(\varDelta) = \varDelta(\int_0^{x_1}(1-\omega_\Theta) - 2\lambda x_1 + \frac{3}{2}\lambda\varDelta)$. For $x_1 \in [0, \varDelta/2[$ this quantity is obviously positive, and for $x_1 \in [\varDelta/2, \varDelta[$, we have $g(\varDelta) \geq \varDelta(\frac{1}{2}\lambda\varDelta - 2\lambda x_1 + \frac{3}{2}\lambda\varDelta) = 2\lambda\varDelta(\varDelta - x_1) \geq 0$. This concludes this point.

This also concludes the proof in the case $n \geq k+2$.

Let us now check the case where $n \leq k+1$. In that case, the only possibility is that $A$ is only composed of one block ($B = A$) and thus $n \in \{k-1, k, k+1\}$. The iterative scheme we built in the previous case ($n \geq k+2$) therefore does not apply here. Our proof will consist of showing that such a set $A$ verifies inequality (6).

Firstly, the case $n = k-1$ implies a configuration where the block intersects both the upper and lower values of $W_{ext}$. The same applies for $W$. Thus $\omega_A(x) = \omega_\Theta(x)$ and $\eta_A(x) - \omega_A(x) = 1 - \omega_\Theta(x), \forall x \in W$. Then $\int_W (1-\omega_\Theta)\int_W \omega_A = \int_W \omega_\Theta \int_W(\eta_A - \omega_A)$ and (6) are clearly verified, and therefore Eq. (5).

Secondly, the case $n = k$ implies that the block intersects only one of the lower and upper values of $W_{ext}$. Let us suppose, without any loss of generality, that the block intersects the upper value of $W_{ext}$. If $x_1 \in [\varDelta/2, \varDelta[$, the proof of the previous case applies.

If $x_1 \in [0, \varDelta/2[$, we have $\int_{\underline{x}-x_1}^{\underline{x}} \eta_A = \int_{\underline{x}-x_1}^{\underline{x}} \omega_A = \int_0^{x_1}(1-\omega_\Theta)$ and we get: $\int_{\overline{x}}^{\overline{x}+x_2} \omega_A = \int_0^{x_2}\omega_\Theta$ and $\int_{\overline{x}}^{\overline{x}+x_2}(\eta_A - \omega_A) = \int_0^{x_2}(1-\omega_\Theta)$, since $\forall t \in [\overline{x}, \overline{x}+x_2]$, $\eta_A(t) = 1$. Then, the computation shows that condition (6) is equivalent to

$$\left(k\lambda\varDelta - \left\{\int_0^{x_1}(1-\omega_\Theta) + \int_0^{x_2}(1-\omega_\Theta)\right\}\right)\int_0^{x_1}\omega_\Theta + \left(\lambda\varDelta - \int_0^{x_1}(1-\omega_\Theta)\right)(k\varDelta - (x_1+x_2)) \geq 0,$$

which is obvious, since $k \geq 2$.

Thirdly, the case $n = k+1$ implies that the block is included in $W_{ext}$. If $x_1 \in [\varDelta/2, \varDelta[$ and $x_2 \in [\varDelta/2, \varDelta[$, the proof of the previous case applies.

If $x_1 \in [0, \varDelta/2[$ and $x_2 \in [\varDelta/2, \varDelta[$, we have $\int_{\underline{x}-x_1}^{\underline{x}} \eta_A = \int_{\underline{x}-x_1}^{\underline{x}} \omega_A = \int_0^{x_1}(1-\omega_\Theta)$ and $\int_{\overline{x}}^{\overline{x}+x_2} \eta_A = \int_{\overline{x}}^{\overline{x}+x_2} \omega_A = \int_0^{x_2}\omega_\Theta$. Then, after simplification, (6) is equivalent to

$$\lambda\varDelta^2(k+1) + x_2\int_0^{x_1}(1-\omega_\Theta) + \int_0^{x_1}(1-\omega_\Theta)\int_0^{x_2}(1-\omega_\Theta)$$
$$+(x_1+x_2)(k-1)\lambda\varDelta + ((k+1)\varDelta - x_1)\int_0^{x_2}(1-\omega_\Theta) \geq 0,$$

which is clearly verified.

If $x_1 \in [0, \Delta/2[$ and $x_2 \in [0, \Delta/2[$, we have $\int_{\underline{x}-x_1}^{\underline{x}} \eta_A = \int_{\underline{x}-x_1}^{\underline{x}} \omega_A = \int_0^{x_1}(1-\omega_\Theta)$ and $\int_{\overline{x}}^{\overline{x}+x_2} \eta_A = \int_{\overline{x}}^{\overline{x}+x_2} \omega_A = \int_0^{x_2}(1-\omega_\Theta)$. Then, after simplification, (6) is equivalent to

$$\lambda\Delta^2(k+1) + \int_0^{x_1}(1-\omega_\Theta)\left(\int_0^{x_2}(1-\omega_\Theta) - k\Delta - (k+1)\lambda\Delta\right)$$
$$+ \int_0^{x_2}(1-\omega_\Theta)\left(\int_0^{x_1}(1-\omega_\Theta) - k\Delta - (k+1)\lambda\Delta\right) + (x_1+x_2)(k-1)\lambda\Delta \geq 0.$$

A study of this function in the same manner as in (a)(i) and (ii) proves that this inequality is verified. Since the proof is identically conducted, for simplicity it is not included in the paper.

This concludes the proof. □

## 4. Conclusion and discussion

The imprecise pignistic transfer method we use to transfer the counts of each cell of the histogram to any interval of the considered reference interval $\Omega$ aims at measuring the effect of granulation on the reconstructed probability. This method is based on the capacity defined by expression (4). Such a transfer method can be used in numerous other applications to achieve interpolations of sampled data (e.g. digital signal processing [14], fuzzy transform [8]) or using granulated-valued histograms to compute continuous indices [4]. The measure of the granulation impact on the reconstructed (interpolated) continuous data it provides via the imprecision of the reconstructed value can be highly helpful for avoiding wrong data interpretations due to granulation in automatic decision processes. The current version of the transfer method only concerns intervals of $\Omega$. It can be seen as a limitation and thus extending the proof to other subsets of $\Omega$ is a very important follow-up of this work. However, since most statistical tools based on histograms use punctual or at most interval estimations of the probability distribution underlying a set of observations, this proof is sufficient to justify several future extensions of these tools.

## References

[1] A. Dame, E. Marchand, Optimal detection and tracking of feature points using mutual information, in: ICIP 2009, Cairo, Egypt, 2009, pp. 3601–3604.
[2] D. Denneberg, Non-additive Measure and Integral, Kluwer Academic Publishers, 1994.
[3] O. Duda, P. Hart, D. Stork, Pattern Classification, second ed., Wiley, 2001.
[4] A. Kuijper, Mutual information aspects of scale space images, Pattern Recognition 37 (12) (2004) 2361–2373.
[5] K. Loquin, O. Strauss, Fuzzy histograms and density estimation, in: SMPS 2006, Third International Workshop on Soft Methods in Probability and Statistics, 2006, pp. 45–52.
[6] K. Loquin, O. Strauss, Histogram density estimators based upon a fuzzy Partition, Statistics and Probability Letters 78 (13) (2008) 1863–1868.
[7] K. Loquin, O. Strauss, On the granularity of summative kernels, Fuzzy Sets and Systems 159 (15) (2008) 1952–1972.
[8] I. Perfilieva, Fuzzy transforms: theory and applications, Fuzzy Sets and Systems 157 (8) (2006) 993–1023.
[9] E. Ruspini, New experimental results in fuzzy clustering, Information Sciences 6 (1973) 273–284.
[10] D.W. Scott, On optimal and data-based histograms, Biometrika 66 (3) (1979) 605–610.
[11] B. Silverman, Density estimation for statistics and data Analysis, Monographs on Statistics and Applied Probability, London: Chapman and Hall, 1986.
[12] J. Simonoff, Smoothing Methods in Statistics, Springer, New York, 1996.
[13] Olivier Strauss, Quasi-continuous histograms, Fuzzy Sets and Systems 160 (17) (2009) 2442–2465, doi:10.1016/j.fss.2009.01.013.
[14] M. Unser, A. Aldroubi, M. Eden, B-spline signal processing: Part i—theory, IEEE Transactions on Signal Processing 41 (1993) 821–833.
[15] T. Wonnacott, R. Wonnacott, Introductory Statistics, Wiley, 1969.
[16] L. Zadeh, Probability measures of fuzzy events, Journal of Mathematical Analysis and Applications 23 (1968) 421–427.